

DISK ARRAY DEVICE

Patent Number: JP9244932
 Publication date: 1997-09-19
 Inventor(s): YAMAMOTO YOSHIHARU
 Applicant(s):: EKUSHINGU:KK; BROTHER IND LTD
 Requested Patent: ☐ JP9244932
 Application Number: JP19960049095 19960306
 Priority Number(s):
 IPC Classification: G06F12/00 ; G06F12/00 ; G06F3/06 ; G06F3/06 ; G11B20/18 ; G11B20/18
 EC Classification:
 Equivalents:

Abstract

PROBLEM TO BE SOLVED: To provide a disk array device capable of reading a target file without reading file management information by an external main computer, shortening the time until the start of the transfer of a file and preventing the efficiency of parallel file transfer from being lowered.

SOLUTION: The file management information stored in a disk device 20 is stored inside the respective disk devices 20 or inside a different storage device. In this case, when an access request for which the identification information of the file is specified is present from the main computer 30, a disk array controller 10 extracts the address information of the specified file based on the file management information and accesses the file present at the address.

Data supplied from the esp@cenet database - I2

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-244932

(43) 公開日 平成9年(1997)9月19日

(51) Int.Cl. ⁶	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 12/00	5 2 0		G 0 6 F 12/00	5 2 0 P
	5 1 4			5 1 4
3/06	3 0 1		3/06	3 0 1 J
	5 4 0			5 4 0
G 1 1 B 20/18	5 7 0		G 1 1 B 20/18	5 7 0 Z

審査請求 未請求 請求項の数 2 O L (全 7 頁) 最終頁に続く

(21) 出願番号 特願平8-49095

(22) 出願日 平成8年(1996)3月6日

(71) 出願人 396004833

株式会社エクシング

名古屋市中区錦3丁目10番33号

(71) 出願人 000005267

ブラザー工業株式会社

愛知県名古屋市瑞穂区苗代町15番1号

(72) 発明者 山本 佳治

愛知県名古屋市昭和区桜山町6丁目104番

地 株式会社エクシング内

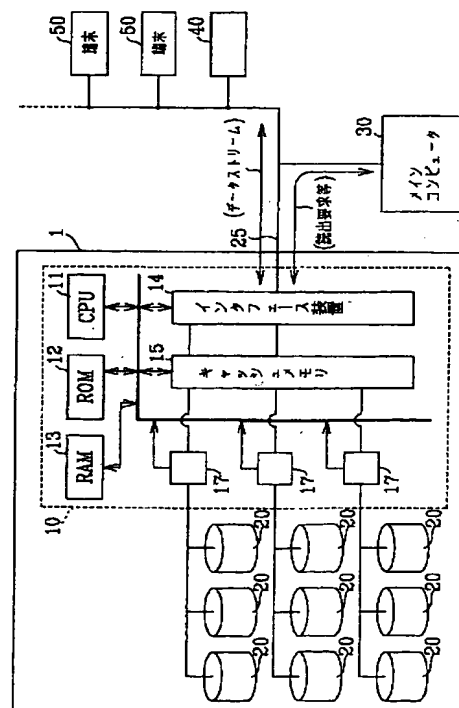
(74) 代理人 弁理士 足立 勉

(54) 【発明の名称】 ディスクアレイ装置

(57) 【要約】

【課題】 ファイル管理情報を外部のメインコンピュータが読み出さないでも目的のファイル読み出しができるようにし、ファイルの転送開始までの時間の短縮や並行するファイル転送の効率を低下させないようにすることのできるディスクアレイ装置を提供する。

【解決手段】 ディスク装置20に格納されているファイル管理情報が、各ディスク装置20内あるいは別の記憶装置内に格納されており、メインコンピュータ30からファイルの識別情報を指定したアクセス要求があった場合、ディスクアレイコントローラ10がファイル管理情報に基づいて指定ファイルのアドレス情報を抽出し、そのアドレスにあるファイルへアクセスする。



【特許請求の範囲】

【請求項1】 ディスクアレイコントローラによって複数台のディスク装置を制御し、メインコンピュータからのアクセスに対しては見かけ上1台のディスク装置として応答するよう構成されているディスクアレイ装置において、

前記ディスク装置に格納されているファイルの少なくとも識別情報及びアドレス情報を含むファイル管理情報が、前記各ディスク装置内あるいは別の記憶装置内に格納されており、

前記ディスクアレイコントローラは、前記メインコンピュータからファイルの識別情報を指定したアクセス要求があった場合には、前記ファイル管理情報に基づいて指定ファイルのアドレス情報を抽出し、その指定ファイルへアクセス可能に構成されていることを特徴とするディスクアレイ装置。

【請求項2】 前記ディスクアレイコントローラは、前記メインコンピュータからのアクセス要求に応じてディスク装置内から読み出したファイルの外部への転送あるいは外部から転送されるファイルのディスク装置内への書き込みを実行している際に、別のファイルに対するアクセス要求に対しても並行して対応可能に構成されていることを特徴とする請求項1に記載のディスクアレイ装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、複数のディスク装置を並列に動作させることで、メインコンピュータからの要求に対して見かけ上1台のディスク装置に見せかけて応答するようにしたディスクアレイ装置に関するものである。

【0002】

【従来の技術】 従来より、複数のディスク装置を並列に動作させることで、メインコンピュータからの要求に対して見かけ上1台のディスク装置に見せかけて応答するようにしたディスクアレイ装置が知られており、このディスクアレイ装置を利用するファイルシステムも考えられている。ディスクアレイ装置を一つの論理ディスクとして使用し、論理ディスクの記憶領域を構成する論理ブロックについて、連続する論理ブロックがそれぞれ別のディスク装置の物理ブロックに割り当てられるように設定することにより、メインコンピュータからのファイルのアクセスに対する高速応答を図るディスクアレイアクセス方法がある。これはストライピングと呼ばれ、RAID (Redundant Arrays of Inexpensive Disks) のレベル4、5等に应用されている。

【0003】 このストライピングを実行すると、ファイルを書き込む際、ファイルを構成するブロック単位のファイルが論理ディスク上の連続したブロックに順次配置される場合には、その連続するブロックはそれぞれ別の

ディスク装置のブロックに配置される。そのため、ファイルをシーケンシャルに読み出す場合には、同一のディスク装置から連続して読み出されることがなく、あるディスク装置のブロックからのデータ転送中に別のディスク装置のブロックをシークしたり、さらにはキャッシュへの先読みなどを行える。このように複数のディスク装置を並列に動作させることができるので、全体として非常に高速なファイル読み出しが可能であった。

【0004】

【発明が解決しようとする課題】 しかしながら、このように1連のファイルを構成するブロック単位のファイルが複数のディスク装置に分散（もちろん上記ストライピングには限定されない）して格納されるような場合には、外部のメインコンピュータがディスクアレイ装置内のデータ読み出しを指示する際、まずディスクアレイ装置内のファイル管理情報を読み出し、ファイルの位置（アドレス）をメインコンピュータから直接指定することにより読み出しを行う必要があった。そのため、このファイル管理情報の読み出しがファイルへのアクセスに先立って必要であり、ファイルの転送開始までに時間がかかっていた。また、複数のファイルを並行して転送するシステムに適用した場合には、時間当たりの転送量を減少させるという問題があった。

【0005】 本発明は、上述した問題点を解決するためになされたものであり、ファイル管理情報を外部のメインコンピュータが読み出さなくても目的のファイル読み出しができるようにし、ファイルの転送開始までの時間の短縮や並行するファイル転送の効率を低下させないようにすることのできるディスクアレイ装置を提供することを目的とする。

【0006】

【課題を解決するための手段及び発明の効果】 この目的を達成するためになされた請求項1記載の発明は、ディスクアレイコントローラによって複数台のディスク装置を制御し、メインコンピュータからのアクセスに対しては見かけ上1台のディスク装置として応答するよう構成されているディスクアレイ装置において、前記ディスク装置に格納されているファイルの少なくとも識別情報及びアドレス情報を含むファイル管理情報が、前記各ディスク装置内あるいは別の記憶装置内に格納されており、前記ディスクアレイコントローラは、前記メインコンピュータからファイルの識別情報を指定したアクセス要求があった場合には、前記ファイル管理情報に基づいて指定ファイルのアドレス情報を抽出し、その指定ファイルへアクセス可能に構成されていることを特徴とするディスクアレイ装置である。

【0007】 本発明のディスクアレイ装置では、ディスク装置に格納されているファイルの少なくとも識別情報及びアドレス情報を含むファイル管理情報が、各ディスク装置内あるいは別の記憶装置内に格納されているので

あるが、ディスクアレイコントローラは、メインコンピュータからファイルの識別情報を指定したアクセス要求があった場合には、ファイル管理情報に基づいて指定ファイルのアドレス情報を抽出し、そのアドレスにあるファイル、すなわち指定されたファイルへアクセスすることができる。

【0008】つまり、ファイル管理情報を外部のメインコンピュータが読み出さないでも目的のファイルを読み出すことができ、ファイルの転送開始までの時間の短縮や並行するファイル転送の効率を低下させないようにすることができる。例えば、メインコンピュータから見れば、ディスクアレイ装置に対してファイルの読み出しを指示すれば、ディスクアレイ装置においてファイル管理情報を読み出して目的のファイル読み出しまで実行するので、その処理が完了した時点で目的のファイルの転送を開始することが可能となる。従来は、まずファイル管理情報をメインコンピュータが読み出し、そのファイル管理情報に基づいてファイルのアドレスを直接指定して読み出しを指示していたが、本ディスクアレイ装置では、そのファイル管理情報を読み出して時間分だけ転送開始を早めることができる。

【0009】なお、ファイル管理情報は少なくとも識別情報及びアドレス情報を含むものであり、その他にファイルサイズや記憶させた日付け等も含むのが一般的である。そして、このファイル管理情報は上述したようにファイル自体を記憶している各ディスク装置に格納されていてもよいし、あるいはディスク装置とは別の記憶装置に格納させてもよい。これはディスク装置と同様にアレイ形式で設けられた例えばICメモリ等でもよいし、ディスクアレイコントローラ内に設けた内部メモリ等でもよい。

【0010】また、請求項2に示すように、ディスクアレイコントローラが、メインコンピュータからのアクセス要求に応じてディスク装置内から読み出したファイルの外部への転送あるいは外部から転送されるファイルのディスク装置内への書き込みを実行している際に、別のファイルに対するアクセス要求に対しても並行して対応可能に構成されているディスクアレイ装置であれば、並行するファイル転送の効率を低下させないようにすることができる。そのため、特にビデオサーバのように、複数の供給先に対して並列にデータを供給するようなシステムに適用した場合には、転送効率を低下しない点の効果は大きい。

【0011】

【発明の実施の形態】次に、本発明の実施形態について図面を参照して説明する。図1は、MPEGビデオデータを扱うファイルサーバにおける大容量の記憶装置としてディスクアレイ装置を適用した実施形態の概略構成を示すブロック図である。

【0012】ファイルサーバは、ディスクアレイ装置1

とそのディスクアレイ装置1に対してデータの読み出し・書き込み等の指示を与えるメインコンピュータ30とから構成されており、両者はバス25を介して接続されている。ディスクアレイ装置1へはバス25を介して外部からMPEGデータを入力したり、ディスクアレイ装置1からバス25を介して外部へMPEGデータを出力することができる。具体的には、例えばMPEGエンコーダ40からのデータが入力されたり、端末50へMPEGデータが出力され、その端末50のMPEGデコーダにおいてビデオ信号としてデコードされることとなる。

【0013】前記ディスクアレイ装置1は、ディスクアレイコントローラ10とビデオサーバとして供給する大容量のビデオデータを記憶しておく複数のディスク装置20とを備えている。このディスク装置20は、いわゆる物理的なハードディスクドライブとそれを制御するコントロールボードとが一体化されたものである。

【0014】ディスクアレイコントローラ10は、制御手段としてのCPU11と、アレイ管理ソフトウェア等の動作プログラムを記憶している記憶手段としてのROM12と、CPU11のワークエリアであるRAM13と、外部とのインタフェースを取るためにバス25と接続されているインタフェース装置14と、キャッシュメモリ15と、プロトコル制御装置17等を備えている。

【0015】キャッシュメモリ15は、ディスク装置20から外部へ出力されるデータ及び外部からディスク装置20へ入力されるデータのディスクキャッシュとして動作する。また、プロトコル制御装置17は、ディスク装置20とキャッシュメモリ15間のデータ転送を行う。図1においては、キャッシュメモリ15に3つのプロトコル制御装置17が接続されており、1つのプロトコル制御装置17についてそれぞれ3台のディスク装置20が接続されているが、必要に応じて、1つのキャッシュメモリ15に4台あるいは8台のプロトコル制御装置17を接続してもよいし、また、1つのプロトコル制御装置17に8台あるいは16台といったディスク装置20を接続してもよい。

【0016】ディスクアレイコントローラ10のCPU11は、ROM12内のアレイ管理ソフトウェア等に基づき、プロトコル制御装置17を介して各ディスク装置20を並列に同時動作させることができる。次に、一例としてDOS (Disk Operating System) と呼ばれているシステムのファイルの取り扱い方法について説明する。通常、図2に示すようにファイルはファイルのアドレス情報であるFAT (File Allocation Table) と、ファイルの情報を記録したディレクトリと呼ばれるファイル管理情報を持っている。本実施形態のディスクアレイ装置1においては、上記FATとディレクトリを含めたファイル管理情報を1台のディスク装置20にまとめて記録されており、ファイルそのものはその他の各ディ

ディスク装置20に記録されている。

【0017】ここでファイル管理情報についてさらに説明する。ファイル管理情報とファイルはクラスタ単位で記録されている。クラスタとは、図3に示されるように磁気ディスクの面を一定の大きさに分割したものであり、データ保存の際の基本的な分割サイズとなる。通常、1、2あるいは4キロバイトである。

【0018】また、ディレクトリには、図4に示すようにファイルの識別情報としてのファイル名、アドレス情報としてのファイルの最初のディスク番号及び最初のクラスタ番号、そしてファイルサイズ等、そのファイルに関する管理情報が格納されている。

【0019】一方、FATは図5に示すようにディスク番号毎に設けられている。各ディスク番号毎に、それぞれ存在するクラスタ分のテーブルがあり、その個々のテーブルは存在するクラスタに対応し、テーブル中にはそのクラスタに連続する次のクラスタエントリと、そのクラスタが存在するディスク番号が格納されている。なお、そのクラスタがファイルの終了クラスタである場合は、例えばFFFF等、他の数値と識別可能な値が格納される。

【0020】次に、ファイルを検索する場合の流れを説明する。まず、fileAを検索する場合には、図4のディレクトリよりディスク番号0のクラスタ番号0000が指定されているので、ディスク番号0；クラスタ番号0000にアクセスし、最初のデータとする。続いて図5に示すディスク番号0でクラスタ番号0000に対応するFATを参照すると、そこにはディスク番号0；クラスタ番号0001が次のクラスタエントリとして示されている。したがって、そのディスク番号0；クラスタ番号0001にアクセスし、次のデータとする。図5においてディスク番号0でクラスタ番号0001に相当するFATにはFFFFという最後のクラスタであることを示す情報が格納されているので、そこで検索処理を終了する。結局ここではディスク番号0；クラスタ番号0000とディスク番号0；クラスタ番号0001のデータをfileAの内容とする。

【0021】同様にfileBに関して検索する場合には、まずディスク番号1；クラスタ番号0000、次にディスク番号1；クラスタ番号0001、続いてディスク番号0；クラスタ番号0002、そして最後にディスク番号1；クラスタ番号0003の順にアクセスして終了する。

【0022】次に、ディスクアレイ装置1からファイルを読み出す場合の動作について説明する。図6はメインコンピュータ30における処理を示すフローチャートである。図6の最初のステップS10においては、ファイル名を指定したオープンコマンドをディスクアレイ装置1に対し発行する。続いて、S20でファイルリードコマンドをディスクアレイ装置1に対し発行する。

【0023】一方、図7はファイルを出力する場合のディスクアレイ装置1での処理、詳しくはCPU11が実行する処理を示すフローチャートである。まず、最初のステップS110において、インタフェース装置14を介してメインコンピュータ30からのファイルオープンコマンドを受ける。すると、続くS120では、ファイル管理情報が格納されている所定のディスク装置20から指定されたファイルのディレクトリエントリを読み出す。

【0024】そして、S130でメインコンピュータ30からファイルリードコマンドを受けると、S140へ移行し、上記S120で読み出したディレクトリエントリから、最初のディスク番号及び最初のクラスタ番号を抽出する。続くS150では、その抽出されたディスク番号のディスク装置20の前記抽出されたクラスタ番号から該当するファイルを読み出す。そして、S160にて、次のクラスタのエントリ番号の指定があるかどうかを判断し、エントリ番号が指定されていたら（S160：YES）、S150へ戻り、その指定されたクラスタ番号にアクセスして該当するファイルを読み出す。このようにS150、S160の処理を繰り返し、最後のクラスタ番号が読み出されるまで該当ファイルの読み出しを行なう。

【0025】そして、最後のクラスタ番号の該当ファイルが読み出されると（S160：NO）、S170へ移行し、その読み出されたファイルをバス25を介して該当する供給先へ出力する。このように本ディスクアレイ装置1では、ディスク装置20に格納されているファイルのファイル管理情報が、所定の1台のディスク装置20に格納されており、ディスクアレイコントローラ10は、メインコンピュータ30からファイルの識別情報を指定したアクセス要求があった場合には、ファイル管理情報に基づいて指定ファイルのアドレス情報を抽出し、そのアドレスにあるファイル、すなわち指定されたファイルへアクセスすることができる。

【0026】つまり、ファイル管理情報を外部のメインコンピュータ30が読み出さなくても目的のファイルを読み出すことができ、ファイルの転送開始までの時間の短縮や並行するファイル転送の効率を低下させないようにすることができる。例えば、メインコンピュータ30から見れば、ディスクアレイ装置1に対してファイルの読み出しを指示すれば、ディスクアレイ装置1においてファイル管理情報を読み出して目的のファイル読み出しまで実行するので、その処理が完了した時点で目的のファイルの転送を開始することが可能となる。

【0027】従来は、まずファイル管理情報をメインコンピュータ30が読み出し、そのファイル管理情報に基づいてファイルのアドレスを直接指定して読み出しを指示していたが、本実施形態のディスクアレイ装置1によれば、必要なデータを読み出すための制御にバス25を

占有する時間が、ファイルオープンとファイルリードのみでよく、バス25への負荷が少なくなる。したがって、そのファイル管理情報を読み出していた時間分だけ転送開始を早めることができる。

【0028】また、ディスクアレイコントローラ10が、メインコンピュータ30からのアクセス要求に応じてディスク装置20内から読み出したファイルの外部への転送あるいは外部から転送されるファイルのディスク装置20内への書き込みを実行している際に、別のファイルに対するアクセス要求に対しても並行して対応可能に構成されている場合には、並行するファイル転送の効率を低下させないようにすることができる。そのため、端末50が多数接続されており、その多数の端末に対して並行してビデオデータを供給する場合には、上述した転送効率を低下しないことによる効果は大きい。

【0029】以上、本発明はこのような実施形態に何等限定されるものではなく、本発明の主旨を逸脱しない範囲において種々なる形態で実施し得る。例えば上記実施形態では、ファイル管理情報を1台のディスク装置20にまとめて記憶させていたが、このファイル管理情報はファイル自体を記憶している各ディスク装置20に格納されていてもよい。但し、この場合にはディスクアレイコントローラ10が実行する図7のS120に相当する処理の際に、各ディスク装置20からファイル管理情報を読み出さなくてはならないので、ディスク装置20の数が多くなった場合には、本実施形態のように1台のディスク装置20にファイル管理情報をまとめて記憶しておく方が便利であり。またデータ読み出しの高速化にもつながる。なお、ディスク装置20ではなく別の記憶装置に格納させてもよい。これはディスク装置20と同

様にアレイ形式で設けられた例えばICメモリ等でもよいし、ディスクアレイコントローラ10内に別途設けたメモリ等でもよい。

【図面の簡単な説明】

【図1】 一実施形態であるディスクアレイ装置をファイルサーバにおける大容量の記憶装置として用いた場合の概略構成を示すブロック図である。

【図2】 ファイル管理情報の概略説明図である。

【図3】 クラスタの分割方法を示す説明図である。

【図4】 ファイル管理情報の内のディレクトリの説明図である。

【図5】 ファイル管理情報の内のFATの説明図である。

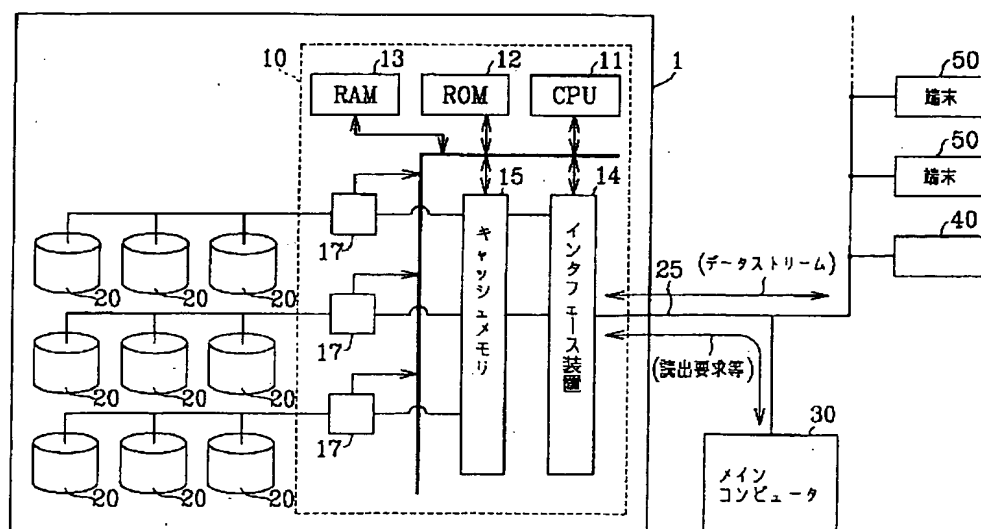
【図6】 ファイルを読み出す場合のメインコンピュータでの処理を示すフローチャートである。

【図7】 ファイルを読み出す場合のディスクアレイコントローラでの処理を示すフローチャートである。

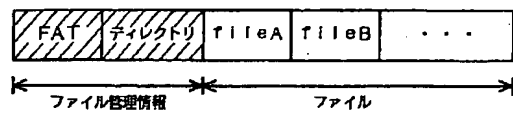
【符号の説明】

1…ディスクアレイ装置	10…ディスクアレイコントローラ
11…CPU	12…ROM
13…RAM	14…インタフェース装置
15…キャッシュメモリ	17…プロトコル制御装置
20…ディスク装置	25…バス
30…メインコンピュータ	40…MPEGエンコーダ
50…端末	

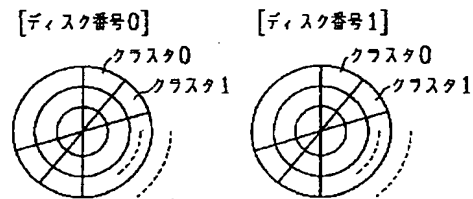
【図1】



【図2】



【図3】



【図4】

【ディレクトリ】

ファイル名		最初の ディスク番号	最初の クラスター番号	ファイル サイズ
fileA	...	0	0000	2Kバイト
fileB	...	1	0000	4Kバイト
...
fileN	...	#	####	###バイト

【図5】

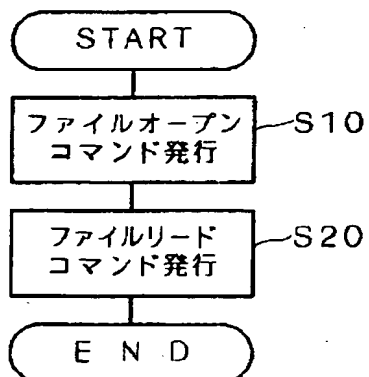
【ディスク番号0のFAT】

対応する クラスター 番号	ディスク 番号	次のクラスター エントリ
0000	0	0001
0001	F	FFFF
0002	1	0003
0003	1	001A
0004	2	0002
0005	0	0007
0006	1	00BC

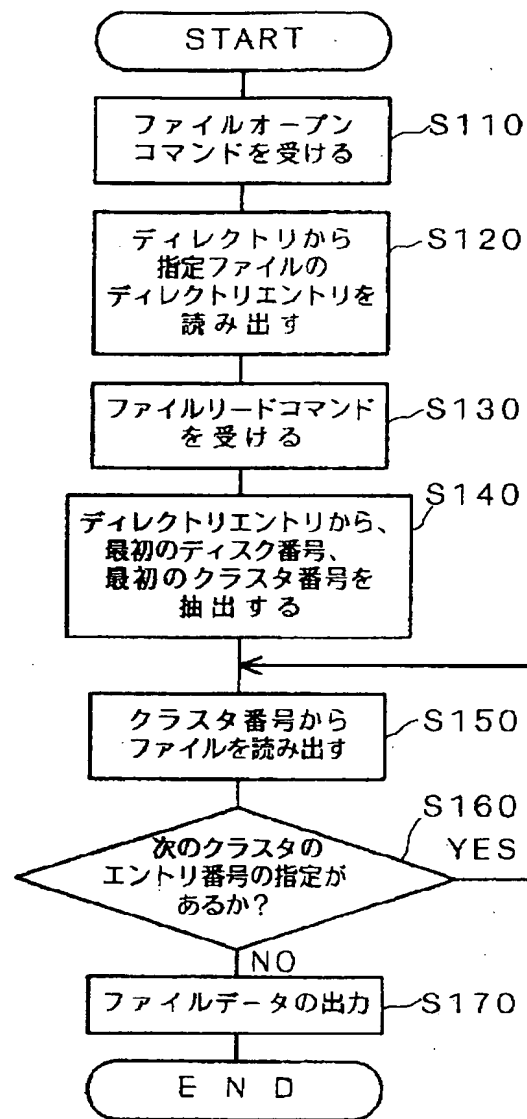
【ディスク番号1のFAT】

対応する クラスター 番号	ディスク 番号	次のクラスター エントリ
0000	1	0001
0001	0	0002
0002	0	00AF
0003	F	FFFF
0004	1	0015
0005	1	0022
0006	0	000B

【図6】



【図7】



フロントページの続き

(51) Int. Cl.⁶
G 1 1 B 20/18

識別記号
5 7 2

庁内整理番号

F I
G 1 1 B 20/18

技術表示箇所

5 7 2 F

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.